

BAB I

PENDAHULUAN

A. Latar Belakang

Perkembangan teknologi pada saat ini dirasakan cepat dan menuntut adanya perubahan berbagai sector perusahaan. Penguasaan ilmu pengetahuan dan teknologi mutlak dibutuhkan guna menunjang perubahan dan perkembangan tersebut. Akibat pesatnya perkembangan tersebut, sebagian besar kehidupan manusia tidak lepas dari pemanfaatan ilmu pengetahuan dan teknologi. Hampir semua bidang membutuhkan untuk penanganan informasi yang mereka miliki, terutama untuk bidang akuntansi sehingga mendorong perusahaan semakin kompetitif.

Menurut Prayitno dalam Ilyas (2001) teknologi adalah perangkat ide, metode, teknik benda-benda material yang digunakan dalam waktu dan tempat tertentu maupun untuk memenuhi kebutuhan manusia.

Dalam beberapa tahun terakhir ini, analisis bisnis dengan menggunakan *data science* (Ilmu data) memiliki peran penting dalam menentukan kebijakan strategi bisnis khususnya meningkatkan pendapatan. Setiap industri dapat mengambil manfaat dari keputusan berbasis data, terstruktur dengan baik. Industri penjualan secara aktif menerapkan solusi ilmu data untuk keuntungannya (N. Calixto and J. Ferreira, 2020).

Inovasi yang dibawa oleh Ilmu Data ke penjualan sebagian besar meningkatkan pengalaman pelanggan dan sebagai hasilnya meningkatkan penjualan. Tentu saja, untuk mencapai tujuan ini banyak pengumpulan data, pemrosesan, pembersihan akan dibutuhkan (F Riza, 2021).

Berdasarkan data dari Kustodian Sentral Efek Indonesia dan PT Raya Saham Registra per 31 Desember 2020 jumlah lembar saham yang beredar sejumlah 123.275.050.000 lembar saham. Sebanyak 35,89% merupakan lembar saham yang dimiliki pemodal asing dari jumlah keseluruhan 45,06% . Pada awal tahun 2019 harga saham sudah mulai melambung berkisar 20,58% yang diikuti oleh aksi beli investor asing (*net sell*) dipasar reguler yang mencapai Rp 2,20 triliun. Harga sahamnya terus menunjukkan tren positif dengan dengan naik terus menerus. Namun pada hari Senin, 04 februari 2019 pemodal asing melepas saham Bank BCA (BBCA). Perbaikan

data ketenagakerjaan Amerika Serikat (AS) membuat para *fund manager* asing menarik dana dari saham-saham berkapitalisasi besar di Indonesia.

Jika di akumulasikan *net sell* (jual bersih) investor asing pada saham BCA mencapai Rp 300,88 miliar dari total nilai transaksi senilai Rp 703,33 miliar. Harga saham BCA terkoreksi 2,34% ke level harga Rp 27.500/saham. Kuatnya data tenaga kerja AS yang sempat memantik aksi beli di pasar saham tanah air pada akhirnya justru menjadi boomerang. Mengutip situs resmi CMA Group yang merupakan pengelola bursa derivative terkemuka di dunia, berdasarkan harga kontrak per 3 februari 2019 bahwa the Fed Fund Futures akan menaikkan suku bunga acuan sebanyak 1 kali (25 basis point) pada tahun 2019 sebesar 4,2% membuat saham-saham perbankan lainnya mengalami tekanan yang cukup dalam (CNBC Indonesia).

Pendekatan ilmu data memberi peluang besar untuk menggunakan data history dan mengubahnya menjadi wawasan yang berguna untuk meningkatkan pendapatan. Mengetahui prediksi penjualan masa depan membawa kelegaan besar bagi perusahaan yang bekerja dengan penjualan. Dengan mengetahui prediksi penjualan masa depan memungkinkan untuk menghindari masalah ini dan membuat keputusan yang lebih baik. (A.M Husein, 2019)

Selain itu, asumsi penjualan keadaan yang berubah yang dapat mempengaruhi penjualan secara signifikan harus ditentukan sebelumnya. Algoritme perkiraan penjualan mencari pola dalam data ini. Pola yang terdeteksi selanjutnya digunakan untuk menilai kecenderungan umum dari transaksi dalam alur untuk membangun prediksi dengan tingkat akurasi yang tinggi. Akan tetapi, hasil evaluasi kinerja algoritma machine learning memberikan hasil yang perlu di pertimbangkan untuk membangun model peramalan yang kuat, selain itu algoritma ini juga dapat diterapkan untuk peramalan kuantitatif maupun kualitatif. (M. Seyedan and F. Mafakheri, 2020).

Machine Learning (ML) adalah aplikasi kecerdasan buatan yang membuat sistem memiliki kemampuan belajar yang berjalan secara otomatis kemudian dapat meningkatkan kemampuannya berdasarkan pengalaman tanpa deprogram secara eksplisit (Kusuma, 2020 dalam Ricky Lie Jaya, 2022). Menurut ilustrasi yang terdapat pada situs ibm.com, ML merupakan salah satu sub bagian dari Artificial Intelligence (AI), pendekatannya berupa

mempelajari pola dari data yang kemudian diingat dalam memori dalam bentuk model, dan model tersebut digunakan untuk memprediksi data yang baru. Oleh karena itu untuk membuat model machine learning diperlukan dataset yang dijadikan proses training, serta dataset yang dijadikan contoh bahan uji untuk menghitung seberapa akurat model tersebut. Contoh penerapan ML yang sekarang sudah banyak adalah model pendeteksian penyakit jantung, alat prediksi cuaca, sistem untuk memprediksi jumlah pelanggan dan masih banyak lagi.

Dalam machine learning terdapat banyak tipe yang dikelompokkan kedalam tiga kelompok, yaitu Supervised Learning, Unsupervised Learning dan Reinforcement Learning. Supervised Learning merupakan algoritma machine learning yang cara kerjanya dengan mempelajari pola data historis untuk memprediksi data yang baru. Unsupervised Learning adalah algoritma yang mempelajari pola dataset untuk mengklasifikasikan ke beberapa kelompok yang memiliki kesamaan karakteristik, unsupervised learning cocok dalam membuat model untuk menentukan segmentasi pasar. (Thomsett, 1998) yang dikutip dari (Button & Barto, 2018).

Penelitian ini fokus membaca data penjualan asing (*foreign sell*) dengan menggunakan kerangka kerja ilmu data untuk memahami dan menemukan wawasan yang berharga dari kumpulan data transaksi penjualan sehingga menghasilkan fitur yang optimal diterapkan pada algoritma perkiraan penjualan. Selanjutnya, hasil analisis data penjualan ini akan diterapkan pada model *machine learning* untuk menemukan algoritma yang paling akurat dalam membaca prediksi penjualan saham. Penelitian ini akan mencari model algoritma *machine learning* untuk memprediksi tren pergerakan harga saham, terutama akan berfokus pada model *Regression* dengan algoritma yang digunakan *Extreme Gradient Boosting(XGBoost) Pipeline4*. berdasarkan penelitian yang dilakukan oleh (Beno Jage, 2022) algoritma XGBoost yang memanfaatkan proses konvolusi menunjukkan hasil yang lebih akurat dari pada algoritma lainnya dalam mengelola data *sequential*, secara arsitektur yang dibentuk lebih sederhana dan jelas. Sedangkan, model *regression* menyediakan fungsi yang menggambarkan hubungan antara satu atau lebih variabel independen dan variabel respon dependen atau target.

Untuk mengatasi permasalahan diatas perlu adanya pengembangan teknologi baru yang lebih efektif dan efisien yaitu dengan membangun

kecerdasan teknologi yang lebih dikenal dengan istilah *Machine learning* yang didukung perangkat keras seperti komputer dan fasilitas internet yang digunakan sebagai perangkat penunjang.

IBM Watson *machine Learning* merupakan platform yang disediakan oleh IBM Watson untuk mendukung pengembangan, implementasi dan manajemen model machine learning. Platform ini dirancang untuk membangun penelitian dan mendeploy machine learning secara efisien. Machine learning sendiri merupakan bagian dari kecerdasan buatan, yang mempelajari pengenalan pola untuk mendapatkan solusi yang optimal. *Machine learning* memungkinkan komputer menemukan solusi data yang optimal secara otomatis (Internasional Business Machine, 2023)

Pembelajaran mesin adalah komponen penting dalam bidang ilmu data yang sedang berkembang. Melalui penggunaan metode statistik, algoritma dilatih untuk membuat klasifikasi atau prediksi, dan untuk mengungkap wawasan penting dalam proyek pengembangan data. Wawasan ini kemudian mendorong pengambilan keputusan dalam aplikasi dan bisnis, yang idealnya berdampak pada metrik pertumbuhan utama. Ketika data besar terus berkembang dan berkembang, permintaan pasar terhadap ilmuwan data akan meningkat. Mereka akan diminta untuk membantu mengidentifikasi pertanyaan bisnis yang paling relevan dan data untuk menjawabnya. (Internasional Business Machines, 2023)

Model machine learning adalah representasi matematis dari suatu sistem atau proses yang menggunakan algoritma dan data untuk "belajar" dari pengalaman (data pelatihan) dan meningkatkan kinerjanya dalam memprediksi atau mengambil keputusan pada data baru yang belum pernah dilihat sebelumnya (data uji). Model ini merupakan inti dari pendekatan *machine learning*, yang bertujuan untuk mengidentifikasi pola dan hubungan dalam data untuk melakukan tugas-tugas seperti klasifikasi, regresi, pengelompokan, atau pengenalan pola (DQLab, 2023)

Model machine learning dapat bersifat sederhana atau sangat kompleks, tergantung pada jenis masalah yang dihadapi dan jumlah fitur yang digunakan. Dalam beberapa kasus, model yang lebih kompleks mungkin memberikan performa yang lebih baik, tetapi dapat menimbulkan risiko overfitting jika tidak diatur dengan baik. Kualitas dan akurasi model sangat bergantung pada kualitas data pelatihan dan proses pengembangan model

yang cermat. Ada beberapa kriteria yang perlu diperhatikan ketika kamu ingin membuat model machine learning yang fit dan juga cocok untuk diimplementasikan (DQLab, 2023).

Berdasarkan latar belakang diatas maka penulis tertarik untuk membuat penelitian memanfaatkan sebuah teknologi kecerdasan buatan untuk membaca sebuah data penjualan asing saham dengan sistem watsonx *auto artificial intelegence machine learning*. Dengan demikian, penelitian ini dapat bermanfaat secara teoritis juga memiliki dampak langsung dalam membantu perusahaan untuk beradaptasi dengan perubahan dalam mencapai pasar saham. Bertilik tolak dari hal tersebut maka penulis tertarik melakukan penelitian dengan judul **“Implementasi Model Machine Learning Dalam Membaca Data Foreign Sell Pada Saham PT Bank Central Asia Tbk. Periode 2019 – 2022”**

B. Rumusan Masalah

Berdasarkan latar berakang diatas maka permasalahan yang dapat diambil adalah “bagaimanakah cara mengimplementasikan model machine learning dengan model Regresi algoritma XGBoost sebagai sarana dalam membaca suatu data penjualan asing pada saham PT Bank Central Asia Tbk ? “

C. Tujuan Penelitian

Tujuan dari penelitian ini adalah membangun visualisasi tren pergerakan saham yang di prediksi oleh model regresi algoritma XGBoost dalam *machine learning* dan mengetahui apakah yang dilakukan dengan model algoritma dapat digunakan sebagai sarana dalam membaca suatu data penjualan asing pada saham PT Bank Central Asia Tbk.

D. Kegunaan Penelitian

Dalam penelitian ini diharapkan akan memberikan manfaat baik secara praktis maupun akademis, sebagai berikut:

1. Kegunaan praktis

Secara praktis diharapkan penelitian ini dapat memberikan manfaat diantaranya:

- a. Mendeteksi anomali: algoritma *machine learning* dapat digunakan untuk mendeteksi aktivitas yang tidak biasa atau potensi ancaman dalam system keamanan komputer dan jaringan.
- b. Pengenalan pola: dapat membantu dalam mengidentifikasi pola keamanan dan memberikan tindakan pencegahan lebih efektif.
- c. Keuangan dan perbankan: membantu dalam menilai risiko dengan menganalisis data keuangan
- d. Deteksi kecurangan: dapat mendeteksi aktivitas kecurangan atau transaksi yang mencurigakan pada sektor keuangan.

2. Kegunaan Akademis

secara akademis diharapkan penelitian ini dapat memberikan manfaat di antaranya:

- a. Penelitian dan analisis data: machine learning membantu menganalisis data besar, mengidentifikasi trend dan merumuskan model prediktif.
- b. Permodelan dan simulasi: machine learning dapat digunakan untuk membangun model dan simulasi dalam berbagai disiplin ilmu.
- c. Informasi: memberikan informasi lengkap dan mempermudah dalam membaca data suatu saham.
- d. Pengembangan algoritma deteksi: meningkatkan keakuratan serta efisiensi algoritma deteksi untuk mengidentifikasi produk, pelanggan, serta pola pembelian.

E. Spesifikasi Penelitian

Spesifikasi produk yang diharapkan dalam penelitian ini adalah sebagai berikut:

1. Model tidak overfitting, model machine learning tidak boleh overfitting karena mengakibatkan performa model yang buruk dan tidak dapat diandalkan saat dihadapkan pada data yang belum pernah dilihat sebelumnya (data uji). Overfitting terjadi ketika model terlalu rumit atau kompleks sehingga mampu “menghafal data pelatihan dengan sangat baik, termasuk noise atau ketidaktepatan yang ada dalam data tersebut. Model yang overfitting memiliki performa yang sangat baik pada data pelatihan, tetapi performa sangat buruk pada data uji. Ini mengakibatkan

model gagal dalam memahami pola yang lebih umum dalam data dan hanya mampu menangkap pola khusus dalam data pelatihan.

2. Memiliki akurasi yang tinggi, machine learning perlu memiliki akurasi yang tinggi karena akurasi adalah ukuran penting untuk menilai sejauh mana model dapat memberikan prediksi atau hasil yang mendekati nilai sebenarnya. Tingkat akurasi yang tinggi menunjukkan bahwa model mampu memahami pola yang ada dalam data dengan baik dan memberikan hasil yang sesuai dengan kenyataan. Model dengan tingkat akurasi yang tinggi lebih dapat dipercaya dan diandalkan untuk memberikan hasil yang benar.
3. Minim tingkat kesalahan, model machine learning harus memiliki tingkat kesalahan yang minim karena tujuan utama dari pengembangan model adalah untuk mencapai performa yang optimal dalam memprediksi atau mengklasifikasikan data yang belum pernah dilihat sebelumnya. Tingkat kesalahan yang minim menunjukkan bahwa model dapat memberikan prediksi yang akurat dan mendekati nilai sebenarnya untuk data uji.
4. Generalisasi yang baik, model machine learning harus memiliki generalisasi yang baik karena generalisasi merupakan kemampuan model untuk menerapkan pengetahuannya dari data yang belum pernah dilihat sebelumnya (data uji) atau situasi dunia nyata yang berbeda. Generalisasi berarti bahwa model dapat mengidentifikasi dan memahami pola yang lebih umum dalam data. Generalisasi yang baik dapat memberikan prediksi yang akurat pada data uji yang belum pernah dilihat sebelumnya.

F. Urgensi Penelitian

Urgensi penelitian berarti seberapa penting pengembangan tersebut dilakukan, seberapa besar pengaruh jika berhasil dalam pengembangan dan seberapa darurat masalah yang timbul jika kita tidak melakukan pengembangan dan mendapatkan solusi dari masalah yang ada (bola.com, 2023)

Pemanfaatan machine learning dalam dunia pendidikan sangat penting karena teknologi ini telah membuka peluang baru untuk meningkatkan efektivitas pembelajaran dan mempersiapkan siswa untuk masa depan yang semakin digital, selain itu juga teknologi sangat bermanfaat bagi pembisnis dan investor (chatgpt, 2024)

Prediksi Pasar dengan Machine Learning salah satu keunggulan utama penggunaan AI dalam investasi adalah kemampuannya untuk membuat prediksi pasar. Dengan mempelajari data historis, AI dapat memberikan perkiraan tentang pergerakan pasar di masa depan. Dengan peran yang semakin dominan dalam dunia investasi, penggunaan teknologi AI membuka peluang baru dan mengubah paradigma di pasar keuangan, investor yang memanfaatkannya dengan bijak memiliki peluang untuk meningkatkan kinerja portofolio mereka (chatgpt,2024)

G. Keterbatasan Penelitian

Tidak hanya kelebihan, machine learning tentu memiliki beberapa kelemahan layaknya teknologi lainnya, berikut beberapa kelemahan adalah:

1. Keterbatasan sumber daya manusia yang menguasai teknologi ini. Teknologi ini termasuk salah satu teknologi yang belum lama booming, sehingga jumlah orang yang benar-benar mahir masih sangat sedikit dibandingkan jumlah kebutuhan yang ada.
2. Rentan terjadi kesalahan, ada banyak hal yang mungkin tidak sesuai dengan harapan dari developer ketika mesin dibiarkan bekerja sendirian untuk menggantikan manusia. Selain itu, untuk proses diagnosis kesalahan yang terjadi ternyata cukup memakan waktu.
3. Membutuhkan waktu agar hasil bisa dilihat. Tentu saja akan terjadu mengingat proses untuk mendapatkan model yang terbaik tidaklah singkat.
4. Perkembangan teknologi machine learning ini dapat menjadi hal yang positif atau negative tergantung dari cara kita menyikapinya. Agar dapat beradaptasi dengan teknologi diperlukannya untuk terus upgrade ilmu.